



Machine learning classification of human joint tissue from diffuse reflectance spectroscopy data

RAJITHA GUNARATNE,^{1,*} ISAAC MONTEATH,¹ JOSHUA GONCALVES,²
RAYMOND SHEH,¹ CHARLES N. IRONSIDE,¹ MICHAEL KAPFER,² RICHARD
CHIPPER,² BRETT ROBERTSON,² RIAZ KHAN,^{2,3,4} AND DANIEL FICK^{2,3}

¹Curtin University, Kent Street, Bentley 6102, Australia

²Australian Institute of Robotic Orthopaedics, 2 Centro Avenue, Subiaco 6008, Australia

³The Joint Studio, 85 Monash Avenue, Nedlands 6009, Australia

⁴Department of Medicine, The University of Notre Dame, Fremantle, Australia

*gunaratne.rajitha@gmail.com

Abstract: *Objective:* To assess if incorporation of DRS sensing into real-time robotic surgery systems has merit. DRS as a technology is relatively simple, cost-effective and provides a non-contact approach to tissue differentiation. *Methods:* Supervised machine learning analysis of diffuse reflectance spectra was performed to classify human joint tissue that was collected from surgical procedures. *Results:* We have used supervised machine learning in the classification of a DRS human joint tissue data set and achieved classification accuracy in excess of 99%. Sensitivity for the various classes were; cartilage 99.7%, subchondral 99.2%, meniscus 100% and cancellous 100%. Full wavelength range is required for maximum classification accuracy. The wavelength resolution must be larger than 8nm. A SNR better than 10:1 was required to achieve a classification accuracy greater than 50%. The 800-900nm wavelength range gave the greatest accuracy amongst those investigated. *Conclusion:* DRS is a viable method for differentiating human joint tissue and has the potential to be incorporated into robotic orthopaedic surgery.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Laser surgery combined with robotic control in orthopaedics is still a developing field which provides the opportunity for more precise surgery, new surgical techniques, and the ability to work remotely with a high level of sterility [1–3]. However, the process of laser surgery does not provide tactile feedback, which is useful for surgeons to determine the type of tissue being ablated and for controlling the ablation depth [4]. Hence there is a risk of iatrogenic damage [5–7].

When light is applied to a tissue type, various wavelengths will be both absorbed and scattered depending on its optical properties [8]. This phenomenon can be measured through a process known as Diffuse Reflectance Spectroscopy (DRS). This study investigates the DRS of human joint tissue to determine the plausibility of automatically distinguishing tissue typing based on the generated spectral data. This was performed using an optical fibre coupled spectrometer with a minimum wavelength resolution of 2.0 nm and Si CCD. It was illuminated using a standard halogen lamp. The resulting data set comprised of spectra from cartilage, subchondral bone, meniscus, and cancellous bone specimens. We also investigated the effects of spectrometer resolution and signal-to-noise ratio (SNR) on classification accuracy.

The objective of this study was to prove with strong statistical significance that the incorporation of DRS sensing into real-time robotic surgery systems has merit. Current orthopaedics surgical technology available in clinical practice relies on specialists (surgeons)

to visually differentiate tissue. DRS as a technology is relatively simple, cost-effective and provides a non-contact approach to tissue differentiation. Its usage as a diagnostic tool has already been proven by identifying various tissue types including bladder [9] where elastic-scatter spectra were obtained using a fibreoptic probe incorporated in a urological cystoscope, brain tissue where near-infrared (NIR) optical-property characterization by measurement of spatially resolved diffuse reflectance [10,11], breast [12,13] tissue where elastic scattering spectroscopy mediated by fibreoptic probes were utilised, cervix [14] tissue which utilised reflectance spectroscopy, colon [15,16] tissue which utilised diffuse reflectance spectroscopy, oesophagus [17,18] tissue utilising fluorescence, reflectance, and light-scattering spectroscopy, ovarian [19] tissue utilising reflectance spectroscopy, pancreas [20,21] tissue utilising optical spectroscopy, and skin [22,23] utilising near-infrared spectroscopy.

2. Method

2.1 Hardware

Human joint tissue in the form of bone and soft tissue specimens were collected from routine total knee replacement surgeries over a 3-month period after clinical ethics and patient consent had been received.

Optical spectra were generated through DRS from a set of 3043 human joint tissue samples which included cartilage, subchondral bone, cancellous bone and meniscus. These spectra had a wavelength range of 200-1000nm and were collected under consistent conditions. The spectrometer (Fig. 1) used for collection was the portable Ocean optics USB-650 tide spectrometer [24] which detected the wavelength dispersed light with a linear silicon CCD array detector and was operated with a wavelength resolution of 2nm. The light source used was a 150W fibre-coupled Halogen lamp, connected to a light ring to standardise illumination spread across the tissue sample. In total, spectra from 1579 cartilage, 1269 subchondral bone, 156 cancellous bone and 39 meniscus samples were collected.

2.2 Sample collection and preparation

Human joint tissue was collected by two orthopaedic surgeons during the course of total knee replacement operations. The tissue was transported to the sensing laboratory, sensed and stored in a freezer within 72 hours of collection.

2.3 Software

Figure 2 illustrates the supervised machine learning work flow used for tissue identification. All computations were performed using the WEKA machine learning tool kit [25] based on the normalised spectra. The data set consisted of 3043 spectra whose measurements spanned across 2048 wavelength channels. Each of these wavelength channel were regarded as an attribute towards the identification of the associated tissue class. There were four tissue classes consisting of cartilage, subchondral bone, cancellous bone, and meniscus.

As part of a supervised learning procedure, the first step involved the identification of the samples to create a ground truth for each sample. This identification was performed by clinical orthopaedic surgeons based on the shape, colour, presentation and, most prominently, the location in which they were removed from the patient.

The second and third steps involved the normalisation and dimensionality reduction of the spectra respectively. Normalisation began through division by the light source spectrum followed by the application of a standard normal variate (SNV) [26] transformation to centre and scale them. The average spectra and standard deviation for each tissue class was then calculated based on this normalised form. This enabled both the inter-class variation and the intra-class variation to be measured. Dimensionality reduction involved reducing the number of attributes or wavelengths associated with each spectral sample. This was achieved through

Multiclass Fisher’s Linear Discriminant Analysis (Multiclass FLDA) [27] and resulted in each sample having only 3 identifying attributes.

The final step was the comparison of present classifiers to determine which best correlated with the ground truth. This was achieved through Linear Discriminant Analysis (LDA) [28] with 10-fold cross-validation being used to determine the resulting classifier accuracy. This involved splitting the data into 10 sets with 9 sets being used for training and 1 set used for testing. This was repeated 10 times with a mean accuracy recorded across the iterations.

The quality and quantity of data required for efficient classification was determined by investigating how the classification accuracy changed depending on the spectral range, resolution and SNR. Changes resulting from the range were determined by running the classifier on 100nm segments of the spectra. This provided a comparative accuracy measure for each wavelength range that can be useful in identifying regions better suited to tissue identification. Changes resulting from the resolution and SNR were determined using a combination of linear-interpolation and noise that was added by artificially degrading the data. The optimal level of noise required was determined by incorporating it within 10 samples. The classifier was then run on this reduced set with the degree at which the results differed from their originals being used as an indicator of how much noise could be tolerated.

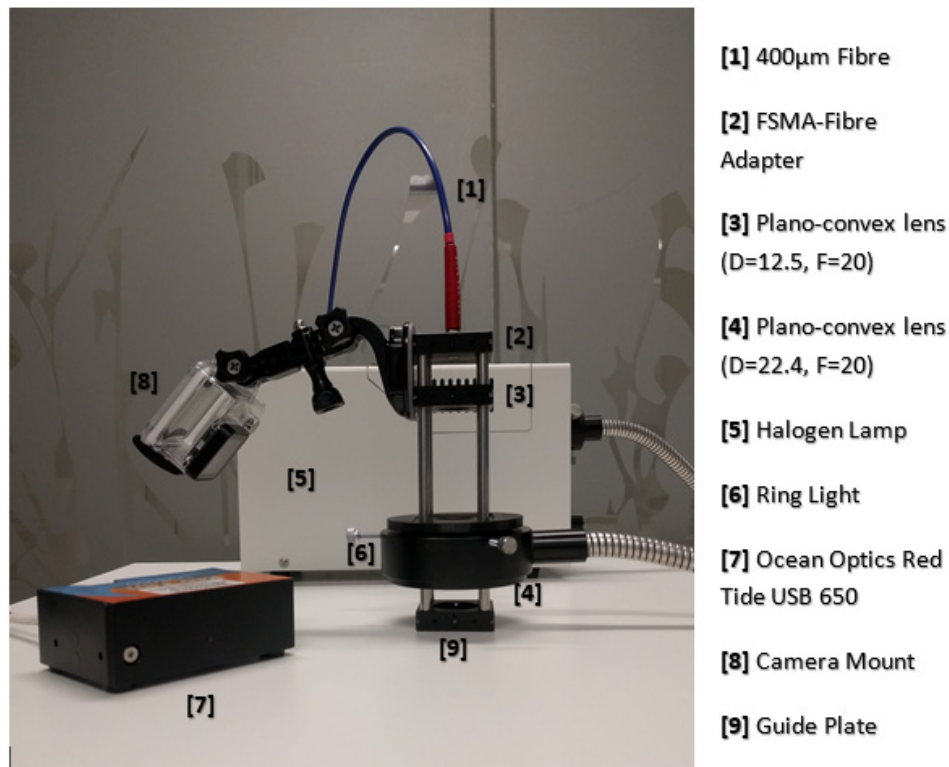


Fig. 1. Spectrometer and optics setup.

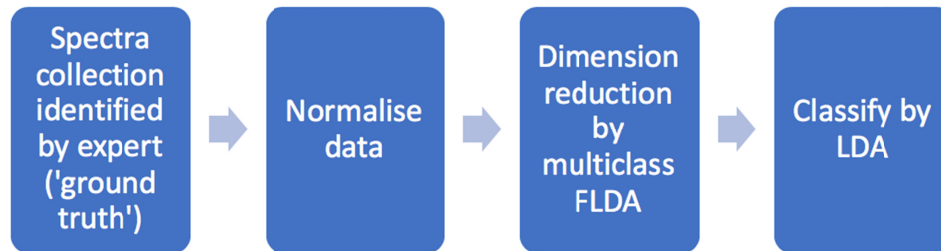


Fig. 2. Machine learning pathway. FLDA: Fisher's linear discriminant analysis, LDA: Linear discriminant analysis.

3. Results

A total of 3034 samples were collected from surgery and normalised. Figure 3 shows the average DRS spectra for each tissue type based on these samples. The number of dimensions or otherwise wavelengths that each contained was then reduced from 2048 to 3 according to their class identifiability. These attributes were the eigenvectors with the largest eigenvalues (24, 25). Many of the WEKA machine learning tools were explored using this data set with the greatest accuracy being achieved with a combination of Multiclass FLDA and the LDA classifier.

Table 1 shows the classification confusion matrix. The classification accuracy was over 99% in 10-fold cross-validation. It is important to note that, even if random data was used, the accuracy would still remain at approximately 50%. This is because the variance between the number of samples from each class is quite significant, with the largest sample consisting of over half the total population. The classifier would therefore simply assign all classes to the largest class which would guarantee it at least this level accuracy. Sensitivity for the various classes were; cartilage 99.7%, subchondral 99.2%, meniscus 100% and cancellous 100%.

Figure 4 shows the classifier accuracy based on different 100nm wavelength ranges. 800-900nm range achieved the highest level of classification accuracy at ~89%. The range 600-900nm comprises the top three wavelength segments for classification accuracy.

There exists a relationship between the size of the wavelength range and the classifier accuracy that results from it. The full wavelength range of 200-1000nm was required to achieve the greater level of accuracy. However, based on the data seen within Fig. 4, an approximate accuracy of 89% can still be achieved using a significantly reduced range of 100nm. Further investigation revealed that classification accuracy is more dependent on the size of a particular wavelength range in comparison to where it actually exists.

Figure 5 shows the classifier accuracy based on wavelength resolution calculated by averaging wavelength ranges whose size increases along the horizontal axis. This demonstrates that the wavelength resolution could be decreased to approximately 8 nm before classifier accuracy drops below 90%.

An SNR ratio of 10:1 was required to degrade classification accuracy to approximately 50%.

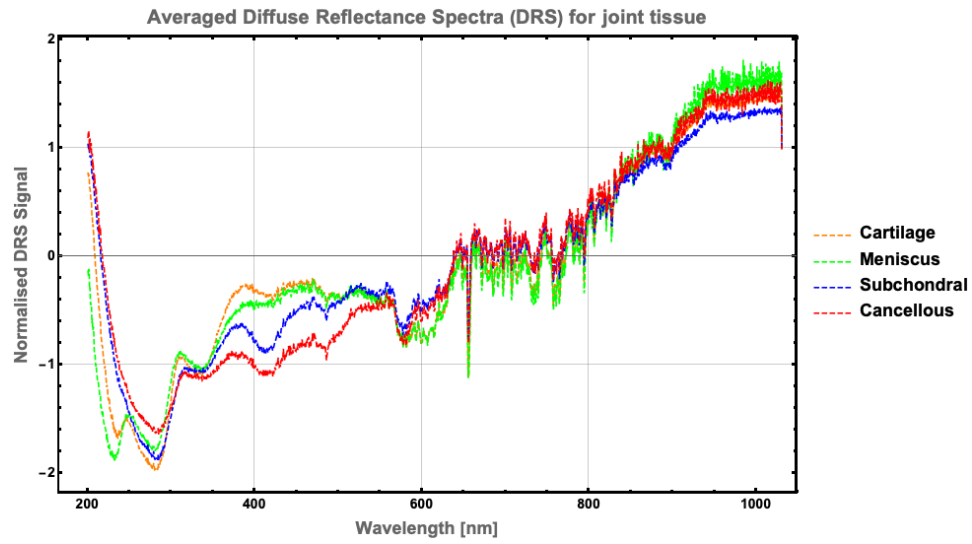


Fig. 3. Average DRS spectra for each tissue type after SNV normalization

Table 1. Machine learning using multiclass FLDA + LDA classifier. A confusion matrix is a table that is used to describe the performance of a classification model. The classification accuracy was >99% in 10-fold cross-validation. Absolute values are in brackets.

		Classified As			
		Cartilage	Subchondral	Meniscus	Cancellous
Class	Cartilage	99% (1574)	1% (5)	0	0
	Subchondral	1% (10)	99% (1259)	0	0
	Meniscus	0	0	100% (39)	0
	Cancellous	0	0	0	100% (156)

Accuracy of Multiclass FLDA + LDA Classifier on Wavelength Ranges

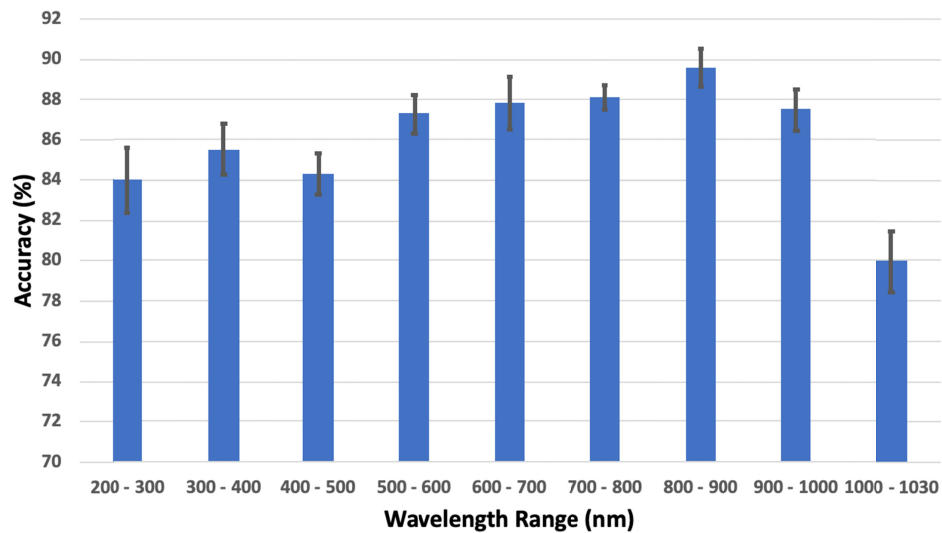


Fig. 4. Accuracy of classifier for different 100nm wavelength ranges with 95% confidence intervals shown as error bars- determined from the 10 fold cross validation procedure.

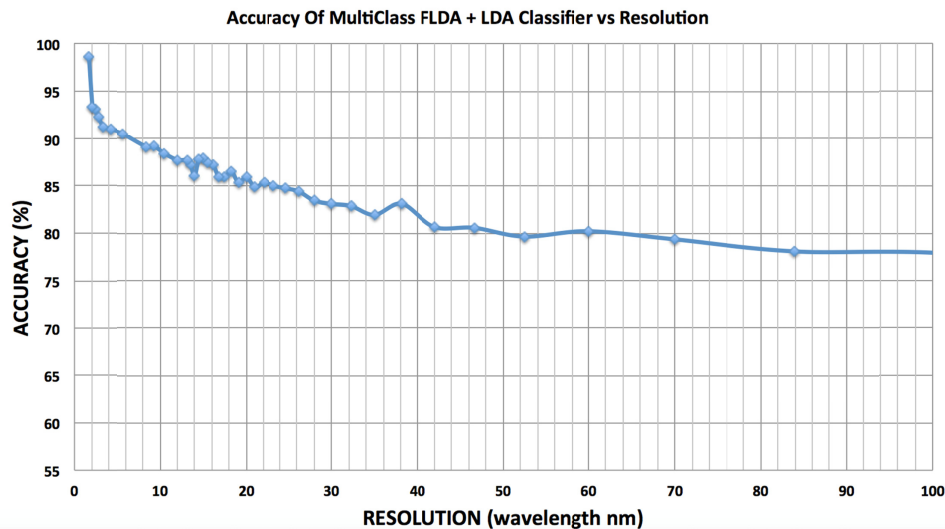


Fig. 5. The classification accuracy versus wavelength resolution- the same machine learning tools were used in each wavelength range multiclass FLDA dimension reduction followed by LDA classification. The wavelength resolution was altered in software by averaging over adjacent wavelengths.

4. Discussion

To achieve the automatic classification of human joint tissue, we employed normalised diffuse reflectance spectroscopy data and WEKA machine learning tools that consisted of Multiclass FLDA for dimensionality reduction followed by LDA classification. This procedure has produced promising results with high classification accuracy. The receiver operating characteristic (ROC) curve passes through the upper left corner reflecting the high overall accuracy. An ROC curve is a plot of the false positive rate (specificity) along the x-axis vs the true positive rate (sensitivity) along the y-axis [29]. Previous optical spectroscopy studies (Raman) have had success with principal component analysis (PCA) for dimensionality reduction [30], however the greatest accuracy we achieved was 88.04% utilizing KNN (K Nearest Neighbours) classifier after PCA.

The fundamental reason for the variation in spectra between different tissue classes is the components within the tissue, such as hemoglobin, lipids, collagen and water. Hemoglobin has an absorption peak in the 400-450nm wavelength range but may also be strongly absorbed in the 200-600nm range. Lipids have an absorption peak in the 900-1000nm range but are also strongly absorbed in ranges above 1000nm. Collagen peaks around the 1500nm range but can similarly be absorbed above the 1000nm range as well. Water is weakly absorbed in visible wavelengths but has increased absorption in both the UV (100-400nm range) and NIR (750-1400nm range) regions (26-28). In our investigation, the range 600-900nm comprises the top three wavelength segments for classification accuracy. This may be due to the high haemoglobin absorption below 600 nm and dominant water and fat absorption above 900nm affecting tissue discrimination.

Reduced classification performance can often arise due to inconsistent collection conditions when collection large quantities of spectral data over an extended period. Differences in global trends, total spectral energy, noise and background of spectra can be caused by inconsistencies in the light source, measurement distance, or the spectrometer [31]. This can make spectra difficult to compare and can cause performance deterioration with classifiers, as spectra of similar materials do not appear similar. It may also cause a classifier to erroneously assign high importance to features that are merely artefacts of the inconsistent conditions. To address this problem, many spectral classification solutions include a

normalisation pre-processing step to adjust for these variations. Common approaches for normalising spectra include constant shifts, smoothing, scaling, SNV, baseline correction and continuum removal [31–36].

It was recognized that a class imbalance was present within this work based on the significant variance existing between class samples sizes. This was investigated to determine if it led to the subsequent positive results. Of all the tissue classes used, Meniscus was the smallest with a data set of 39. The same number of data samples were taken from the remaining classes and the classifier was ran on the new data set. Regardless of this reduction, it still achieved a classification accuracy of above 99%. This demonstrates that suitable discrimination must exist between the classes, independent of any influence from their sample sizing.

Cross-validation is a technique used to evaluate predictive models by partitioning the original sample into a training set to train the model, and a test set to evaluate it [37]. The 10-fold cross-validation is a common form of this where the data set is broken into ten different sets where nine of these are used to compose the training set and the remaining one is used for the test set. The classification accuracy was above 99% using this technique. This accuracy is similar to that achieved by Stelzle et al. [38] where Principle Component Analysis (PCA) and Quadratic Discriminant Analysis (QDA) were used to achieve 94.8% accuracy on 12,150 pig tissue measurements. However, our method learned a comparable level of accuracy from approximately 4 times fewer measurements.

Figure 5 demonstrates that accuracy does not degrade below 90% until resolution drops below approximately 8nm. A colour camera has a spectral resolution of around 70-100nm [39]. This suggests we could achieve around 80% accuracy with a normal colour camera. At a 10:1 signal to noise ratio, the classifier was only able to correctly classify approximately 50% of the time, hence the classifier requires a signal to noise ratio better than 10:1.

5. Conclusion

We have used supervised machine learning in the classification of a DRS human joint tissue data set and achieved classification accuracy in excess of 99%. We employed a halogen light source and a spectrometer for DRS of human joint tissue collected from surgical procedures. We collected a data set comprising diffuse reflectance spectra of 4 types of human joint tissue in the wavelength range 200nm-1030nm. They included 1579 cartilage, 1269 subchondral bone, 156 cancellous bone and 39 meniscus samples. Tests on the data set revealed that full wavelength range is required for maximum classification accuracy, the wavelength resolution must be larger than 8nm, a SNR better than 10:1 was required to achieve a classification accuracy greater than 50% and the 800-900nm wavelength range gave the greatest accuracy amongst those investigated

Disclosures

The authors declare that there are no conflicts of interest related to this article

References

1. J. J. Kuttenger, S. Stübinger, A. Waibel, M. Werner, M. Klasing, M. Ivanenko, P. Hering, B. Von Rechenberg, R. Sader, and H. F. Zeilhofer, "Computer-guided CO₂-laser osteotomy of the sheep tibia: technical prerequisites and first results," *Photomed. Laser Surg.* **26**(2), 129–136 (2008).
2. S. Stopp, D. Svejdar, E. von Kienlin, H. Deppe, and T. C. Lueth, "A new approach for creating defined geometries by navigated laser ablation based on volumetric 3-D data," *IEEE Trans. Biomed. Eng.* **55**(7), 1872–1880 (2008).
3. P. Spinelli, G. Calarco, A. Mancini, and X. G. Ni, "Operative colonoscopy in cancer patients," *Minim. Invasive Ther. Allied Technol.* **15**(6), 339–347 (2006).
4. F. Stelzle, A. Zam, W. Adler, K. Tangemann-Gerk, A. Douplik, E. Nkenke, and M. Schmidt, "Optical nerve detection by diffuse reflectance spectroscopy for feedback controlled oral and maxillofacial laser surgery," *J. Transl. Med.* **9**(1), 20 (2011).
5. G. D. Baxter, D. M. Walsh, J. M. Allen, A. S. Lowe, and A. J. Bell, "Effects of low intensity infrared laser irradiation upon conduction in the human median nerve in vivo," *Exp. Physiol.* **79**(2), 227–234 (1994).

6. T. Menovsky, M. van den Bergh Weerman, and J. F. Beek, "Effect of CO₂ milliwatt laser on peripheral nerves: Part I. A dose-response study," *Microsurgery* **17**(10), 562–567 (1996).
7. T. Menovsky, M. Van Den Bergh Weerman, and J. F. Beek, "Effect of CO₂-Milliwatt laser on peripheral nerves: part II. A histological and functional study," *Microsurgery* **20**(3), 150–155 (2000).
8. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**(11), R37–R61 (2013).
9. J. R. Mourant, I. J. Bigio, J. Boyer, R. L. Conn, T. Johnson, and T. Shimada, "Spectroscopic diagnosis of bladder cancer with elastic light scattering," *Lasers Surg. Med.* **17**(4), 350–357 (1995).
10. F. Bevilacqua, D. Piguet, P. Marquet, J. D. Gross, B. J. Tromberg, and C. Depeursinge, "In vivo local determination of tissue optical properties: applications to human brain," *Appl. Opt.* **38**(22), 4939–4950 (1999).
11. A. N. Yaroslavsky, P. C. Schulze, I. V. Yaroslavsky, R. Schober, F. Ulrich, and H. J. Schwarzmaier, "Optical properties of selected native and coagulated human brain tissues in vitro in the visible and near infrared spectral range," *Phys. Med. Biol.* **47**(12), 2059–2073 (2002).
12. I. J. Bigio, S. G. Bown, G. Briggs, C. Kelley, S. Lakhani, D. Pickard, P. M. Ripley, I. G. Rose, and C. Saunders, "Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results," *J. Biomed. Opt.* **5**(2), 221–228 (2000).
13. G. M. Palmer, C. Zhu, T. M. Breslin, F. Xu, K. W. Gilchrist, and N. Ramanujam, "Monte Carlo-based inverse model for calculating tissue optical properties. Part II: Application to breast cancer diagnosis," *Appl. Opt.* **45**(5), 1072–1078 (2006).
14. D. Arifler, C. MacAulay, M. Follen, and R. Richards-Kortum, "Spatially resolved reflectance spectroscopy for diagnosis of cervical precancer: Monte Carlo modeling and comparison to clinical measurements," *J. Biomed. Opt.* **11**(6), 064027 (2006).
15. G. Zonios, L. T. Perelman, V. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, and M. S. Feld, "Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo," *Appl. Opt.* **38**(31), 6628–6637 (1999).
16. Z. Ge, K. T. Schomacker, and N. S. Nishioka, "Identification of colonic dysplasia and neoplasia by diffuse reflectance spectroscopy and pattern recognition techniques," *Appl. Spectrosc.* **52**(6), 833–839 (1998).
17. M. B. Wallace, L. T. Perelman, V. Backman, J. M. Crawford, M. Fitzmaurice, M. Seiler, K. Badizadegan, S. J. Shields, I. Itzkan, R. R. Dasari, J. Van Dam, and M. S. Feld, "Endoscopic detection of dysplasia in patients with Barrett's esophagus using light-scattering spectroscopy," *Gastroenterology* **119**(3), 677–682 (2000).
18. I. Georgakoudi, B. C. Jacobson, J. Van Dam, V. Backman, M. B. Wallace, M. G. Müller, Q. Zhang, K. Badizadegan, D. Sun, G. A. Thomas, L. T. Perelman, and M. S. Feld, "Fluorescence, reflectance, and light-scattering spectroscopy for evaluating dysplasia in patients with Barrett's esophagus," *Gastroenterology* **120**(7), 1620–1629 (2001).
19. U. Utzinger, M. Brewer, E. Silva, D. Gershenson, R. C. Blast, Jr., M. Follen, and R. Richards-Kortum, "Reflectance spectroscopy for in vivo characterization of ovarian tissue," *Lasers Surg. Med.* **28**(1), 56–66 (2001).
20. R. H. Wilson, M. Chandra, J. Scheiman, D. Simeone, B. McKenna, J. Purdy, and M. A. Mycek, "Optical spectroscopy detects histological hallmarks of pancreatic cancer," *Opt. Express* **17**(20), 17502–17516 (2009).
21. S. Y. Lee, W. R. Lloyd, M. Chandra, R. H. Wilson, B. McKenna, D. Simeone, J. Scheiman, and M. A. Mycek, "Characterizing human pancreatic cancer precursor using quantitative tissue optical spectroscopy," *Biomed. Opt. Express* **4**(12), 2828–2834 (2013).
22. E. Salomatina, B. Jiang, J. Novak, and A. N. Yaroslavsky, "Optical properties of normal and cancerous human skin in the visible and near-infrared spectral range," *J. Biomed. Opt.* **11**(6), 064026 (2006).
23. G. Zonios, A. Dimou, I. Bassukas, D. Galaris, A. Tsolakidis, and E. Kaxiras, "Melanin absorption spectroscopy: new method for noninvasive skin investigation and melanoma detection," *J. Biomed. Opt.* **13**(1), 014017 (2008).
24. O. Optics, "USB-650 red tide spectrometers," <https://oceanoptics.com/product/usb-650-red-tide-spectrometers/>.
25. G. Holmes, A. Donkin, and I. H. Witten, "WEKA: a machine learning workbench," in *Proceedings of second Australia and New Zealand conference on intelligent information systems* (Institute of Electrical and Electronics Engineers, 1994), pp. 357–361.
26. R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc. Rev.* **43**(5), 772–777 (1989).
27. A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, "Linear discriminant analysis: A detailed tutorial," *AI Commun.* **30**(2), 169–190 (2017).
28. J. C. Gonzalez, "Using linear discriminant analysis (lda) for data explore: step-step," <https://www.apsl.net/blog/2017/07/18/using-linear-discriminant-analysis-lda-data-explore-step-step/>.
29. R. Kumar and A. Indrayan, "Receiver operating characteristic (ROC) curve for medical researchers," *Indian Pediatr.* **48**(4), 277–287 (2011).
30. L. F. C. S. Carvalho, M. S. Nogueira, L. P. M. Neto, T. T. Bhattacharjee, and A. A. Martin, "Raman spectral post-processing for oral tissue discrimination - a step for an automatized diagnostic system," *Biomed. Opt. Express* **8**(11), 5218–5227 (2017).
31. T. W. Randolph, "Scale-based normalization of spectral data," *Cancer Biomark.* **2**(3–4), 135–144 (2006).
32. A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.* **36**(8), 1627–1639 (1964).
33. R. Gautam, S. Vanga, F. Ariele, and S. Umapathy, "Review of multidimensional data processing approaches for raman and infrared spectroscopy," *EPJ Tech. Instrum.* **2**(1), 8 (2015).

34. K. H. Liland, A. Kohler, and N. K. Afseth, "Model-based pre-processing in Raman spectroscopy of biological samples," *J. Raman Spectrosc.* **47**(6), 643–650 (2016).
35. R. Kokaly, D. Despain, R. Clark, and K. Livo, "Spectral analysis of absorption features for mapping vegetation cover and microbial communities in yellowstone national park using AVIRIS data," in *Integrated Geoscience Studies in the Greater Yellowstone Area— Volcanic, Tectonic, and Hydrothermal Processes in the Yellowstone Geoecosystem*(2007), pp. 463–489.
36. R. A. Jarvis, "On the identification of the convex hull of a finite set of points in the plane," *Inf Process* **2**(1), 18–21 (1973).
37. M. L. Open, "10-fold Crossvalidation," <https://www.openml.org/a/estimation-procedures/1>.
38. F. Stelzle, C. Knipfer, W. Adler, M. Rohde, N. Oetter, E. Nkenke, M. Schmidt, and K. Tangermann-Gerk, "Tissue discrimination by uncorrected autofluorescence spectra: a proof-of-principle study for tissue-specific laser surgery," *Sensors (Basel)* **13**(10), 13717–13731 (2013).
39. C. Matasaru, "Mobile phone camera possibilities for spectral imaging," in *School of Computing Department* (University of Eastern Finland, 2014).